Hsin-Ling (Justin) Hsu

Google Scholar: Hsin-Ling Hsu
GitHub: JustinHsu1019
LinkedIn: justinhsu101999

justin-code.com
justin.hsu.1019@gmail.com

Education

National Chengchi University (NCCU) — Expected Graduation: Jun. 2027
B.S. in Management Information Systems (MIS) — *Taipei, Taiwan*

- **Rank: 2 / 38 | GPA: 4.3 / 4.3**
- Previously enrolled in Mathematical Sciences (Sept. 2023 – Aug. 2024), and have since continued collaborating with Prof. Jengnan Tzeng from the department on research projects [1] and [3].
- **Technical Courses:** Calculus, Linear Algebra, Statistics(I), Machine Learning, Operations Research/Management Science, Computer Programming, Data Structure, Database Management Systems, Introduction to Computer Science, Introduction to Data Analysis and Programming, Python for Data Analysis 101, An Introduction to Game Theory (I)
- **Related Domain Courses:** The Brain and I, AI and Society

Research Interests

Information Retrieval; AI for Healthcare; Large Language Models; Natural Language Processing

Publications†

[1] **Hsin-Ling Hsu** (first author), Ping-Sheng Lin, Jing-Di Lin, and Jengnan Tzeng. *"KAP: MLLM-assisted OCR Text Enhancement for Hybrid Retrieval in Chinese Non-Narrative Documents."* Peer-reviewed research paper accepted to the AI CUP Workshop @ NTCIR-18, to appear in Proceedings of the 18th **NTCIR** Conference (indexed by DBLP), Tokyo, Japan, 2025. *(paper link)* *(GitHub link)*

[2] **Hsin-Ling Hsu*** (first author), Cong-Tinh Dao*, Luning Wang, Zitao Shuai, Thao Nguyen Minh Phan, Jun-En Ding, Chun-Chieh Liao, Pengfei Hu, Xiaoxue Han, Chih-Ho Hsu, Dongsheng Luo, Wen-Chih Peng, Feng Liu, Fang-Ming Hung, and Chenwei Wu. *"MedPlan: A Two-Stage RAG-Based System for Personalized Medical Plan Generation."* Accepted to **ACL** 2025 Industry Track (~25.1% acceptance rate, Avg. Rating: 7.00 [6–8]), Vienna, Austria, 2025. *(paper link)*

** Equal contribution.*

† All listed publications are peer-reviewed and presented at international conferences.

On-Going Work

[3] **Hsin-Ling Hsu** (first author) and Jengnan Tzeng. *"DAT: Dynamic Alpha Tuning for Hybrid Retrieval in Retrieval-Augmented Generation."* Under Review at the 2nd Conference on Language Modeling (**COLM**), Montreal, Canada, 2025. *(paper link)*

Work Experience

Research Assistant — Dec. 2024 – Present
Far Eastern Memorial Hospital — *New Taipei, Taiwan*

- Advisor: Dr. Fang-Ming Hung
- [2] Research focuses on developing and training models for disease prediction, medical plan generation, and brain activation and epileptogenic zone prediction using LLMs, CLIP, CNN, Knowledge Graph, information retrieval techniques. These models leverage electronic health records (EHR) from outpatient, inpatient, and emergency departments, along with examination reports and SOAP notes.

AI Intern — Sept. 2024 – Present
GoFreight (The world's largest cloud-based freight forwarding software) — *Taipei, Taiwan*

- Developed LLM-powered dynamic web parsing solutions to mitigate crawler disruptions caused by web changes, significantly reducing maintenance costs.
- Applied OCR, NLP, and computer vision techniques to extract and analyze logistics documents (e.g., Master Bill of Lading, Invoice), improving the accuracy and efficiency of automated document processing.
- Researched cutting-edge AI developments and conducted internal presentations on topics such as LLM Agents architecture principles and their potential business applications, enhancing the company's AI strategic roadmap.

Research Assistant — May 2024 – Present
Institute of Information Science, Academia Sinica (The highest-level academic institution in Taiwan) — *Taipei, Taiwan*

- Affiliated with Prof. Ti-Rong Wu's Reinforcement Learning and Games Lab, supporting the research project *General Computer Game Solving based on Proof Cost Network (PCN)* through system development.
- Responsible for designing and developing both frontend and backend components of computer games (e.g., Go, Kill-All Go, and other board size variants) using WGo.js, C/C++.
- Designed a distributed multi-database framework for managing large-scale game records and AI-analyzed board states, ensuring system stability, scalability, and efficient data access across components in a high-performance gaming environment.

Part-time Engineer Jul. 2023 – Sept. 2024

ChainSea Information Group *Taipei, Taiwan*

- AI Engineer | Core R&D Contributor to Open-Source LLM and Whisper Projects at ChainSea, specializing in real-time transcription, inference acceleration, dataset generation and augmentation (e.g., Taipower project under Selected Projects), and model training with expertise in fine-tuning and LoRA techniques.
- Designed, developed, and optimized RAG architectures to improve knowledge base retrieval accuracy.
- Conducted R&D in advanced AI technologies through the development of a dialogue system designed to engage individuals struggling with drug addiction, leveraging LLM Agent architectures and a conversational topic management framework based on stack- and branch-based structures integrated with function calling mechanisms.
- Technologies utilized: Python, SentenceTransformer, TensorFlow, PyTorch, Unsloth.

Awards & Honors

2nd Place in HOTAI MaaS Hackathon, [2/233 teams; ~0.8%] 2024

AI Intelligent Travel Checkup. News Article | GitHub | Certificate *Taiwan*

- A nationwide competition open to all ages, hosted by two giant corporations, Hotai Motor and Microsoft.
- By presenting an innovative forum and AI algorithm for intelligent itinerary check-ups and recommendations, we won second place nationwide and **received a prize of 250,000.**
- I was responsible for system architecture design, AI & full-stack development, as well as delivering the technical presentation and demo.
- Tech stack: Vue.js, LLM with Hybrid Retreival and Cross-Encoder for intelligent POI suggestions, multi-platform review analysis, and personalized activity recommendations using collaborative filtering.

3rd Place in LINE FRESH Campus Competition, [3/165 teams; ~1.8%] 2024

AI dementia care platform. News Article | GitHub | Certificate *Taiwan*

- A nationwide competition hosted by LINE.
- I was responsible for system architecture design, AI & backend development.
- Tech stack: LINE Messaging API, RAG architecture with LLM integration for multilingual support, and health tracking system with data synthesis capabilities. Features include AI-powered chatbot for dementia care guidance, personalized health monitoring with automatic reporting, and community-based knowledge sharing platform.

2nd Place in AI Interdisciplinary Sustainability Innovation Competition, [2/44 teams; ~4%] 2024

Campus AI assistant AllPass Project Lead. News Article | Live Demo | Certificate *Taiwan*

- A competition hosted by National Chengchi University (NCCU).
- For more details, please refer to the Selected Projects section below.

Selected Projects

AutoMouser (100+ stars) 2025 – Present

Open Source Contributor | Pull requests | Issues

- AutoMouser leverages LLM-based technology to automatically generate browser automation code from your mouse movements, capturing every click, drag, and hover. This integration streamlines your workflow and enables the creation of robust, repeatable tests with enhanced precision and flexibility.
- Contributing new features, bug fixes, and codebase architecture optimization.

NCCUPass APP 2024 – Present

Role: Head of AI | Advisor: Prof. Augustin Lu

- Led the AI team in the R&D of the campus AI assistant AllPass and Campus Smart Lost and Found Matching FindPass. AllPass utilizes RAG technology with multi-turn conversation memory, while FindPass combines YOLO for object detection with DINOv2 and Text Embedding Models for vector representations and weighted similarity matching.
- As of September 2024, **over 2,000 students** at National Chengchi University have registered and used the platform, and the project has received multiple accolades in startup competitions.

Taiwan Power Intelligent Robot Optimization Project 2023

ChainSea Information Group & Taiwan Power (Taiwan's largest electric power company)

- Served as a **core contributor** in leveraging LLMs to enhance and optimize training data for Taiwan Power's official website intelligent AI customer service system, "Dianbao," with the currently deployed version utilizing the improved data we developed.

Activities

Lecturer Sept. 2023 – Aug. 2024

Google Developer Group (GDG) - NCCU *Taipei, Taiwan*

- Teaching Experience: Introduction to Databases, Generative AI (Course satisfaction rate as high as **90%**)

Security Research 2024 – Present

HITCON ZeroDay Platform *Remote*

- Reported multiple security vulnerabilities, prompting rapid emergency patches from affected organizations.

Freelancer 2023 – Present

Successfully completing 10+ software development and technical consulting projects. *Remote*

- Experienced in web scraping, LLM-based stock fundamental analysis data processing, ML stock price prediction project consulting, AI for board games, etc.